

A Neural Network to Classify Fatigue from Human-Computer Interaction

André Pimenta^{a,*}, Davide Carneiro^a, José Neves^a, Paulo Novais^a

^a *CCTC/Department of Informatics, University of Minho, Braga, Portugal*

Abstract

Fatigue, especially in its mental form, is one of the most worrying health problems nowadays. It affects not only health but also motivation, emotions and feelings and has an impact both at the individual and organizational level. Fatigue monitoring and management assumes thus, in this century, an increased importance, that should be promoted by private organizations and governments alike. While traditional approaches are mostly based on questionnaires, in this paper we present an alternative one that relies on the observation of the individual's interaction with the computer. We show that this interaction changes with the onset of fatigue and that these changes are significant enough to support the training of a neural network that can classify mental fatigue in real time. The main outcome of this work is the development of non-invasive systems for the continuous classification of mental fatigue that can support effective and efficient fatigue management initiatives, especially in the context of desk jobs.

Keywords: Mental Fatigue, Ambient Intelligence, Neural Networks, Human Computer Interaction

2010 MSC: 00-01, 99-00

1. Introduction

In most of the so-called developed countries people have nowadays increasingly busier lifestyles. This makes them stretch their limits to find time for work,

*Corresponding author

Email address: `apimenta@di.uminho.pt` (André Pimenta)

leisure, family and friends. This necessary extra time is frequently obtained at
5 the expense of smaller periods of sleep or rest, and with a cost in terms of pressure and stress. This, together with other issues such as increasing workloads or harsher work conditions, results in the emergence of fatigue as one of the most worrying health conditions of this century [1, 2]. Although the effects of fatigue may not be readily visible, they have consequences at many levels other than
10 health, including emotions, productivity, performance, social behaviour, among others [3].

Generally, the term fatigue is used to describe a series of manifestations that range from drowsiness or lack of concentration to lack of physical strength or agility [4]. Thus, the concept of fatigue is very broad and subjective. It can
15 be seen as a combination of symptoms, including impaired performance (loss of attention, slowed reaction and response times, impaired decision making, and poor performance on tasks that usually reflect good performance), and subjective feelings of sleepiness and tiredness [5, 6]. It can also be divided in two main types [7]: mental fatigue (when it is our cognitive abilities that
20 are decreased) and physical fatigue (when we become physically impaired due to an excess of a specific physical activity). In this paper we focus on mental fatigue, although a relationship between the two types, that sometimes interact, is acknowledged.

Fatigue can be seen all around us, on an everyday basis: students approach-
25 ing an exam that sleep less in order to have more time to study, office or industrial workers working long, night or consecutive shifts, healthcare professionals subject to long hours of highly wearing conditions, drivers of heavy machinery, or specialized personnel such as the military or fire fighters. Either in trivial tasks or in more specialized ones, fatigue is regarded as one of the main causes
30 of human error [2]. It may lead, especially in the aforementioned scenarios, to serious life-or-death situations, when people working in risky jobs have their performance, attention or motivation impaired due to fatigue.

Fatigue can occur at any time during the day and is not necessarily related to the Circadian Rhythm [8] alone but results of more complex set of factors

35 that include time of rest, motivation, physical or cognitive impairments or social
environment [9]. Depending on its duration and intensity, fatigue can make the
carrying out of daily tasks increasingly hard or even impossible [10]. In severe
or prolonged cases it can cause illnesses such as depression or chronic fatigue
syndrome, with an impact at a personal, social and economic level. Given all
40 this, it is nowadays clear that the detection, monitoring and management of
fatigue should be considered as a priority in the healthcare domain, with orga-
nizations having a special responsibility in their human resources management
strategies.

This paper deals with the issue of fatigue monitoring, with the aim of pro-
45 viding a non-intrusive, reliable and easy to use tool that can be used freely in
organizations, without changing or interfering with the established work rou-
tines. Specifically, we look at desk jobs and others similar, in which people
spend long hours interacting with the computer.

In preliminary work we have established that our interaction patterns with
50 the computer, measured in terms of the use of the keyboard and mouse, change
when under fatigue [11]. In this paper we take this work a step forward by
presenting a service-based architecture, suited for the collection of data for the
purpose of classifying fatigue. We detail the whole process, from the definition
of the architecture to the real-time collection of data. Finally, we describe the
55 process through which soft computing techniques are used to classify fatigue,
in real-time, from the interaction with the computer. Specifically, we train a
Neural Network that takes as input features that describe the user’s interaction
patterns and provides as output an estimation of the user’s mental fatigue.

The paper is structured as follows. Following this introductory section we
60 analyse related work in what concerns the analysis and detection of fatigue,
both using subjective and objective measures. As we do this, we analyse exist-
ing approaches and compare them with the one proposed in this paper. The
paper then moves on to describe the service-based architecture developed to
implement the collection of data. The following section is devoted to a specific
65 case-study, in a real environment, concerning the detection of fatigue. It de-

tails the process of data collection and analysis. Finally, the paper ends with the analysis and discussion of the results attained with this work, and with the concluding remarks and analysis of future directions for this line of research.

2. Related Work

70 From the existing literature it is possible to conclude that there are three main types of approach to the problem of mental fatigue detection and monitoring. The first relies in quantifications of cognitive performance, often measured as the combination of the speed and success or accuracy of an individual when carrying out a task. The second approach consists on subjective ratings of
75 mental workload and mental fatigue. Finally, physiological indicators can also reveal subtle but precise aspects of mental fatigue that are often not revealed by performance measures or subjective ratings.

2.1. Performance Measures

Task performance may be defined as the relative competence of the indi-
80 vidual's interpretation of task-associated stimuli, and the effectiveness of his responses on the task. Excessive mental and physical workload and very low workload (observed in monotonous tasks) are frequent scenarios in which performance decrements are observed. Perhaps the most widely used indicator of fatigue is the degradation of task performance over time or in response to
85 changes in task demands. Welford [12] distinguished four types of performance changes that may occur, namely:

- Impairment of sensory or perceptual functions;
- Slowing of sensory-motor performance;
- Irregularity of time; and
- 90 • The disruption of performance.

Performance decrement due to fatigue may be considered as the total quantity and/or quality of work performed as a function of time on task [6]. Cognitive performance is often measured as the combination of the accuracy per response time of an individual when carrying out a task. This combination is often called
95 throughput [13]. Response speed is the inverse of response time and is more normally distributed than response time in data collected from fatigued operators. Both speed and accuracy are usually acceptable when the individual has normal cognitive skills, and one or the other or both are decreased when these skills are impaired or diminished by fatigue. When mental fatigue takes place, human
100 performance is not consistent, with changes in speed, accuracy and variability [14]. These changes allow us to measure the performance of an individual on a task and are the base of the work proposed in this paper further ahead.

2.2. Subjective Measures

Subjective measures of mental workload are often used to assess the mental
105 workload associated with a task [15, 16]. Obtaining mental workload levels during task performance can be a difficult procedure. However, estimations of subjective workload levels can help isolate task characteristics that affect performance. The two instruments of this type most often used in research were developed in parallel in the 1980s, one at the NASA-Ames Research Center in
110 California and the other in the U.S. Air Force Human Factors Research Group in Wright-Patterson AFB, Ohio.

The NASA Task Load Index (NASA-TLX) is a multidimensional assessment tool [15]. The main seven-point scale is:

115 *Overall Performance: How successful were you in performing the task? How satisfied were you with your performance?*

The TLX has five seven-point sub-scales that help identify difficult task characteristics. The sub-scales are:

- 120 • Mental Demand: How much mental and perceptual activity was required?
Was the task easy or demanding, simple or complex?
- Physical Demand: How much physical activity was required? Was the
task easy or demanding, slack or strenuous?
- Temporal Demand: How much time pressure did you feel due to the pace
125 at which the tasks or task elements occurred? Was the pace slow or rapid?
- Frustration Level: How irritated, stressed, and annoyed versus content,
relaxed, and complacent did you feel during the task?
- Effort: How hard did you have to work (mentally and physically) to ac-
complish your level of performance?

130 The Subjective Workload Assessment Technique (SWAT) uses a three-dimensional
model of workload: time load, mental effort load, and psychological stress load.
There are three rating levels for each dimension: low, medium and high. Usually,
subjects perform a pre-sort of the 27 possible combinations of the dimensions
and their sort is transformed into an interval scale of workload ranging from 0
135 to 100. For the actual task, each activity or event is rated by assigning a value
of 1 to 3 on each of the three dimensions. The pre-sort scale value associated
with this combination becomes the workload value for that activity or event.

A seven-point mental workload scale was created by the Crew Performance
Branch of the USAF School of Aerospace Medicine in the late 1970s, and then
140 re-examined, linearised, and verified by the Human Factors Engineering Branch
of the Air Force Flight Test Center [17]. This is the easiest rating instrument
to use in high workload environments, but it provides less insight than the
TLX or SWAT instruments. Individuals choose one of seven sets of statements
describing their average mental workload during the preceding work period:

- 145 • Nothing to do; no system demands.
- Light activity; minimum demands

- Moderate activity; easily managed; considerable spare time.
- Busy; challenging but manageable; adequate time available.
- Very busy; demanding to manage; barely time enough.
- 150 • Extremely busy; very difficult; non-essential tasks postponed.
- Overloaded; system unmanageable; essential tasks undone; unsafe.

Other tools exist to estimate mental fatigue, namely the Vigor-Activity and Fatigue-Inertia factors of the Profile of Mood States (POMS). The POMS measures five aspects of affect or mood [18, 19]. It consists of 65 adjectives describing
 155 feelings and mood to which the subject responds according to a five-point scale ranging from "Not at all" to "Extremely".

Other similar tools exist, such as the Karolinska Sleepiness Scale (KSS) [20, 21], the Stanford Sleepiness Scale (SSS) [22] or the Pittsburgh Sleep Quality Index (PSQI).

160 2.3. Physiological Measures

Several physiological variables can be used as indicators of fatigue. Electroencephalography (EEG) is a technique that has been used often for the detection of patterns of mental fatigue and also for the analysis of cognitive overload. Alan Gevins has conducted EEG studies to analyse workload and
 165 fatigue [23, 24]. Oculography and eye blink measures such as PERCLOS have been used for fatigue detection [25, 26, 27].

Wrist activity monitors (WAM) have been used to track sleep and sleep debt. The best-validated WAM is the ActiGraph, which uses the Cole-Kripke algorithm [28, 29]. Other well-known physiological measures exist, such as heart
 170 rate (HR), heart rate variability (HRV) and the electrodermal response (EDR), that have however been shown to be somewhat unreliable for the detection of mental fatigue [30, 31].

A relatively new discipline, called Neuroergonomics, has arisen in the domain of physiological measurement [32]. Its goals parallel the goal of the performance-based approach that is the focus of this brief literature review. Neuroergonomics takes advantage of modern tools that provide information about brain functions, including EEG, Magnetoencephalography and Functional Magnetic Resonance Imaging (fMRI).

Parasuraman and Wilson reviewed neuroergonomics studies in workload and vigilance, adaptive automation, neuroengineering, and molecular genetics and individual differences [33]. They concluded that the neuroergonomics approach can enrich our understanding of the use of technology by humans and can inform technological design. They suggested that applications should include the assessment of operator workload and vigilance, the implementation of real-time adaptive automation, neuroengineering for people with disabilities, and design of selection and training methods.

2.4. Critical Analysis of Current Approaches

On the face of so many different approaches to the problem, one must enquire whether it makes sense to propose yet another one. In this exercise let us first recall the aims of our research line: we aim to develop an approach to fatigue detection and classification that can be used continuously and non-intrusively in real-life scenarios such as the workplace.

From this point of view, physiological approaches present a main problem, despite their accuracy: their continuous use in work environments is somewhat impractical, perhaps with the exception of monitoring eye and eyelid movements remotely (although image-based approaches are also frequently declined by workers on the grounds of violating privacy or promoting work surveillance). In general terms these approaches require trained personnel to be present to place sensors and other hardware, which is also often expensive. Taking the EEG as example, it requires the placement of scalp sensors that usually lead to significant limitation on the individual being monitored, especially in what concerns movements, besides the time it takes for an accurate placement. More-

over, there is another caveat, especially in work environments, that must also be stated: some individuals and some workers' unions object physiological measures in the workplace under the grounds of violating privacy. They label them
205 as "medical" measures, despite some of them not possessing medical precision, and claim that they collect information that can be used for many other purposes.

Concerning subjective instruments, their main problem is that they do not
210 fully take into account inter-individual differences. Although there are some that help account for individual differences, such as the Profile of Mood States (POMS) and the State-Trait Anxiety Inventory (STAI) [34, 35], their use in complex systems can prove to be complicated and confusing, due to the same problems that can be observed in subjective measures of fatigue detection [36, 37, 38].
215 Subjective measurements also have the problem of being made through questionnaires that are filled by the user himself, interfering with the primary task, besides from the traditional problems related to wording, question construction and interpretation or answer biasing.

The use of the decrement of performance as an indicator of mental fatigue
220 is also not without criticism, the main issue being the inexactness in accounting for the decrement. Fatigue effects may not be disclosed when subjects compensate by exerting greater effort. Usually, substantial or excessive time on task is required to induce task-specific fatigue or acute fatigue. However, a very high workload can cause this same effect in less time. In both cases there are
225 excessive task demands during the time required to perform a task. Excessive workload conditions can be considered as accelerating the build-up of fatigue, while normal workload results in a slower emergence of fatigue.

Performance decrement is also often not observed, even though performance requirements or workload are increased. This phenomenon may be explained
230 by the concept of spare capacity, a concept developed originally for the assessment and modelling of humans as drivers and pilots [39, 40]. The theory of spare capacity proposes that man rarely works at their maximal work capacity, and can thus temporarily absorb additional cognitive or physical workload while

Table 1: Summary of the strengths and weaknesses of the three main types of approaches.

Type	Strengths	Weaknesses
Performance	-Can be acquired through non-invasive methods -Strongly related to fatigue	-Sensitive to spare capacity effect
Subjective	-Easy and inexpensive to implement -Many validated questionnaires and other tools available	-Based on the individual's subjective perceptions -Interferes with work routine
Physiological	-Very accurate -Impossible to fake	-Expensive and complex approach -Often requires specific hardware and trained personnel -Impractical and intrusive

maintaining a standard level of performance, given that his spare capacity is not
 235 exceeded. Thus, the increase of fatigue is often viewed as decreasing in individual's spare capacity, and the concepts of workload and fatigue are considered to be interconnected.

The strengths and weaknesses pointed out and summarized in table 1 motivate and support the development of the approach proposed in this paper:
 240 one that is non intrusive and can thus be used continuously throughout the day. It will improve access to this kind of initiatives, with a positive impact in individual health, absenteeism and related costs, productivity and quality of work.

3. An Architecture for Distributed Data Acquisition

245 Prior to carrying out the research work aimed at understanding the influence of fatigue on the interaction with the computer, a suitable framework for data

acquisition was idealized and developed. We clearly aimed at a framework in line with the Ambient Intelligence (AmI) philosophy [41], in which the technological component is hidden in the environment and the user is placed in the middle of the paradigm. There is also a focus on non-intrusiveness, with acquisition of information taking place without the need for explicit or conscious user interactions.

AmI systems are commonly described as electronic environments that seamlessly interact and adapt to human needs, in which people are surrounded by intelligent and intuitive interfaces embedded in all kinds of objects. To take full advantage of the information gathered ubiquitously from various sources in the environment there is a need for a software infrastructure that allows an easy integration, promotes interoperability, and focuses on extensibility. Considering these aspects, in this work we present an infrastructure to support an efficient approach for the development of AmI applications, following an approach based on Service Oriented Architectures (SOAs). Indeed, SOAs are being increasingly adopted in both the academic and industrial arenas, even to integrate Multi-agent Systems [42].

The more appropriate way of doing so is to adopt a Service-Component Architecture (SCA): a group of OASIS specifications that has become an industry standard. It is intended for the development of applications based on SOA, which defines how computing entities interact to perform work for each other. Originally published in November 2005, SCA is based on the notion that all the functions in a system should exist in the form of services that are combined into composites to address specific business requirements. In other words, it allows to build service-oriented applications as networks of service components. SCA is used for building service components, assemble components into applications, deploy to (distributed) runtime environments and reuse service components built from new or existing code using SOA principles. This approach is advantageous in AmI for the following reasons:

- Interoperability. Provides loose coupling allowing integration without the

need to know how components are implemented. Components can be written using any language, and can use any communication protocols and infrastructure to link them, making it easier to integrate components to form composite applications;

- Maintainability. Composition of solutions is clearly described as a declarative application of infrastructure services. Simplification for all developers, integrators and application deployers;
- Flexibility of Development. Service Components are easier to develop because the semantics of each independent Service Component are significantly less complex than the overall of a single, (relatively large) monolithic application; each Service Component can be developed by a different team of developers, each of whom focusing only on their component without having to know the details of work done by others. Components can easily be replaced by other components and services can be easily invoked either synchronously or asynchronously;
- Reusability. Since each Service Component has well-defined interfaces, each component can be developed, tested and debugged independently of the other components. This not only speeds up project implementation but, in the case of well-designed Service Components, also leads to significantly enhanced reuse of components;
- Dynamic Deployment and Runtime Modification/Replacement. Service Components can be dynamically deployed to remote nodes at runtime, and components within a process can be easily replaced by new or updated components, further reducing the time taken to modify or change an existing process in response to business requirements;
- Configuration Management and Version Control. Service Components facilitate version control and dynamic configuration management, allowing fine-grained control over deployments across the enterprise.

SCA provides a good basis for AmI applications [43], it is in line with our architectural model and it fulfils major AmI deployment requirements by promoting late bindings at deploy time and runtime with the support of several relevant technologies including POJO, SOAP, REST, BPMN, BPEL, JMS, Camel or Rules services. But most of all it is currently supported by several major commercial and open source products such as Jboss Switchyard, IBM WebSphere or TRENTINO (C++).

From the several available implementations of SCA we have chosen JBoss SwitchYard since it is an open source solution in a relative mature state, and also enhances some of the SCA advantages. Specifically, SwitchYard advocates transparency when running a service during its whole lifecycle. Important aspects such as connectivity, orchestration and routing do exist in SwitchYard in a modular format, which means one can deploy them in an independent way. Using a SwitchYard graphical user interface (Fig. 1), one can build visual models of the applications, that are meant to improve the software engineer's ability to comprehend and communicate the full composition of their applications and also to speed up development and integration projects.

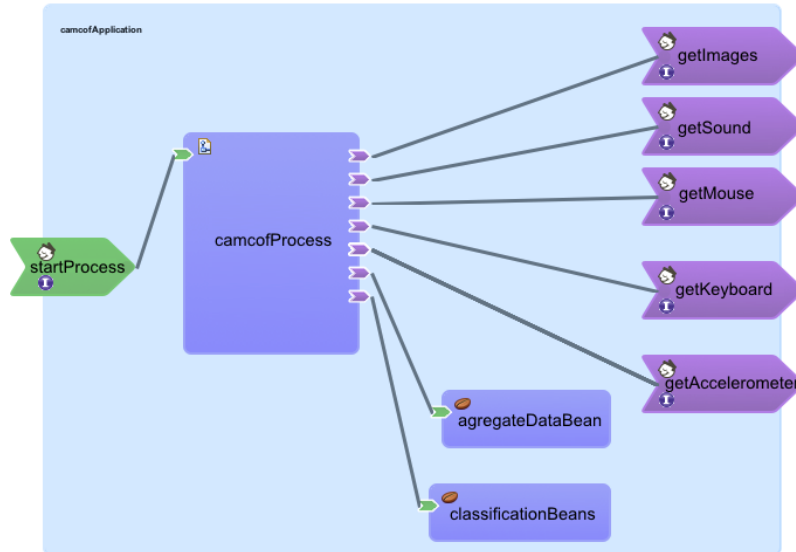


Figure 1: Application composition using SwitchYard graphical user interface.

A service-based approach was followed to develop an architecture logically divided into several packages that encapsulate a set of features and tasks. Figure 2 presents its high-level view, detailing the five packages that compose the system. The Runtime Environment is the main component of the architecture, where system operations are executed. It is also through this environment that the remaining components are accessed. The Service component, executed by the Runtime Environment, contains all services of monitoring and data collection through the use of sensors, and is responsible for encapsulating all external resources collected by the system, represented by the External Resources. The tasks carried out by the Runtime Environment are also supported by queries to components Knowledge Management and Intelligent Component. Knowledge Management supports the management of information and knowledge collected by the system. The Intelligent Component contains components of Intelligent Systems/Artificial Intelligence used in this architecture for data processing.

The Runtime Environment corresponds to the core component of the architecture where the main features are executed. It is also divided in two subcomponents, which are nonetheless interconnected: the Runtime Management and Runtime Platform. The first is a platform for managing processes running on the system. It is composed of a Data Manager that contains an interface to link the Knowledge Manager, an Execution Manager responsible exclusively for the management of the execution of system processes and the Service Directory that contains all information from the execution of services in the architecture. It should also be noted that the Runtime Management has an interface to connect the Service Component. The Runtime Platform is a platform for the execution of the Runtime Environment. This is composed of a Service Bus that, through the interface provided by the Runtime Management, establishes communication between the two subcomponents and also a Service Execution Engine that represents the execution engine services in the architecture.

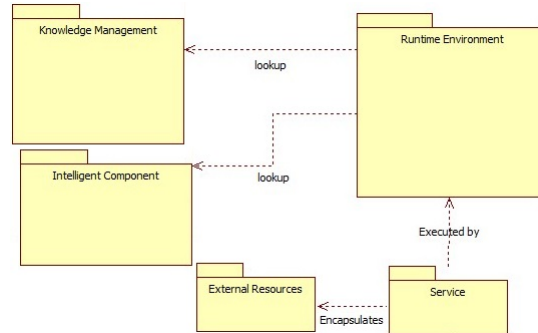


Figure 2: High-level view of the architecture.

350 4. Case-Study on Human Fatigue Detection

An actual implementation of the architecture described in the previous section is currently in place at the Intelligent Systems Lab of the University of Minho. In this environment, a relatively large number of people, mostly students, spend their day or part of it sitting in front of a computer (either their personal laptop or a desktop belonging to the lab), which is constantly connected to the local wireless and/or wired networks. It thus provides the optimum conditions for an ongoing collection of data, that can span several months for each particular individual. The framework can nonetheless be easily installed in other environments since the server, which processes the information received from each computer, is accessible from external locations. This was the case in this particular case-study, as will be described ahead, in which an environment for data collection was set up in a classroom in which students make use of computers.

Although not all this information was used in the current study, the following sources of information are available in the environment:

- **Accelerometers** - Accelerometers placed on chairs, keyboards and mouses, measure how users are moving and how they are using their peripherals;
- **Mouse and Keyboard** - These devices provide information about how the user interacts with the peripherals (e.g. velocity of the mouse, typing

370 rhythm, number of mistakes);

- **Microphone** - Microphones are used to measure the amount of noise in the vicinity of the user, allowing to perceive their social environment;
- **Video Camera** - An estimation of the amount of movement is calculated from the video camera. The image processing is based on difference techniques to calculate the amount of movement between two consecutive frames.

380 All the described sources of information (from now on designated as sensors) are encapsulated by services running locally. Each of these services exposes different features using a Web Service interface. These services were integrated using SwitchYard allowing the service orchestration to support the automation of system processes by loosely coupling services across different applications. There is a clear separation between process logic and Web Services, providing the system with increased flexibility.

385 The main process was modelled using Business Process Modelling Notation (BPMN). In this standard (Fig. 3), a consumer service invokes a process flow via the service interface. The orchestration engine invokes services to process various service which in turn invoke further service requests until the workflow process is completed and results are provided to the service consumer. The service orchestration engine, a component of the SwitchYard, handles the overall process flows, calling the appropriate web services and determining the next steps to complete.

390 With this architecture it is possible to orchestrate the acquisition, transformation, and classification processes used to collect and to extract meaningful information. It is also possible to use different services (encapsulating different and disperse sensors), simultaneously or individually, coordinating all the elements of the main process.

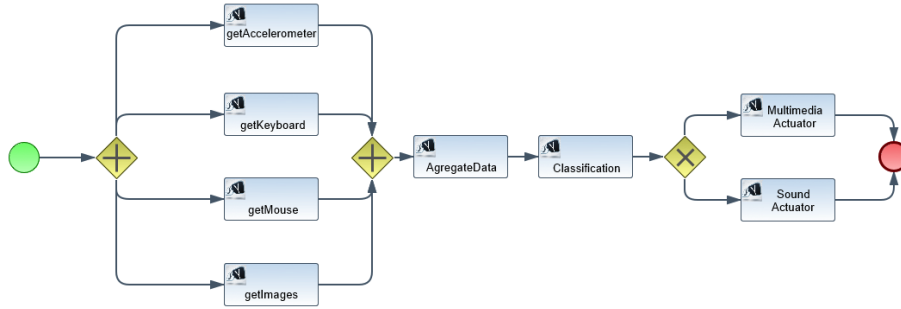


Figure 3: The "camcof" process, viewed through BPMN

4.1. Methodology

The case study detailed in this section was devised in order to achieve the following objectives:

- 400 • To assess the suitability of the architecture to support the necessary functionalities;
- To assess the suitability of the selected features to model the users' interaction patterns;
- To determine if mental fatigue does have a significant effect on the considered features;
- 405 • To assess the possibility of training a Neural Network to classify mental fatigue in real time;

Given these objectives, the main research question of the study is defined as follows: mental fatigue affects the interaction of an individual with the computer, and this can be accurately measured with a non-intrusive approach.

The following methodology was followed to implement the case study. The previously mentioned application for data collection was used during two weeks in two different classes, that take place in different moments of the day. Each session has a duration of three hours. Data is collected continuously and organized on an hourly basis. Each hour of data originates one instance, with each

instance depicting the variance of each feature in that period for a given user, as well as a subjective evaluation of said user’s level of fatigue, provided through a self-report mechanism. Specifically, we used the seven-point USAFSAM Mental Fatigue Scale created by Dr. William F. Storm and Captain (Dr.) Layne P. Perelli of the Crew Performance Branch of the USAF School of Aerospace
420 Medicine, Brooks AFB, San Antonio, Texas, and then used in many field and laboratory tests [44, 6]. This scale classifies fatigue into seven different values, as follows:

1. Fully alert. Wide awake. Extremely peppy.
- 425 2. Very lively. Responsive, but not at peak.
3. Okay. Somewhat fresh.
4. A little tired. Less than fresh.
5. Moderately tired. Let down.
6. Extremely tired. Very difficult to concentrate.
- 430 7. Completely exhausted. Unable to function effectively. Ready to drop

The main objective is to use this user feedback paired with the variance of the feature in each hourly period to train the neural network and thus understand the relationship between the variables.

The participants, twenty four in total (19 men, 5 women) were students
435 from the University of Minho of the field of physical sciences. Their age ranged between 18 and 25. Participants were provided with the application for logging events of the mouse and keyboard. This application, which maintained the confidentiality of the keys used, needed only to be installed in the participant’s primary computer and would run on the background, starting automatically
440 with the Operating System. The only explicit interaction needed from the part of the user was the input of very basic information on the first run, including the identification and age.

4.2. Data Collection

For this particular study, and given the aforementioned objectives, we fo-
445 cused on the data collected from the mouse and the keyboard of the user’s com-

puters. The interaction of the user with the computer is monitored in terms of specific Operating System events fired from the use of the computer's mouse and keyboard:

- MOV, timestamp, posX, posY - an event describing the movement of the mouse, in a given time, to coordinates (posX, posY) in the screen;
- MOUSE_DOWN, timestamp, [Left|Right], posX, posY - this event describes the first half of a click (when the mouse button is pressed down), in a given time. It also describes which of the buttons was pressed (left or right) and the position of the mouse in that instant;
- MOUSE_UP, timestamp, [Left|Right], posX, posY - an event similar to the previous one but describing the second part of the click, when the mouse button is released;
- MOUSE_WHEEL, timestamp, dif - this event describes a mouse wheel scroll of amount dif, in a given time;
- KEY_DOWN, timestamp, key - identifies a given key from the keyboard being pressed down, at a given time;
- KEY_UP, timestamp, key - describes the release of a given key from the keyboard, in a given time;

From these events, that fully describe the interaction of the user with the mouse and keyboard, we extract a set of features, based on notions of behavioural biometrics, that will be used to monitor fatigue:

KEY DOWN TIME (KDT) - the timespan between two consecutive KEY_DOWN and KEY_UP events, i.e., for how long was a given key pressed.

TIME BETWEEN KEYS (TBK) - the timespan between two consecutive KEY_UP and KEY_DOWN events, i.e., how long did the individual took to press another key.

MOUSE VELOCITY (MV) - The distance travelled by the mouse (in pixels)
 475 over the time (in milliseconds). The velocity is computed for each interval defined by two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$, that took place respectively in the instants *time₁* and *time₂*. Let us also assume two vectors *posx*
 480 and *posy*, of size *n*, holding the coordinates of the consecutive MOUSE_MOVE events between *mup* and *mdo*. The velocity between the two clicks is given by $r_dist/(time_2 - time_1)$, in which *r_dist* represents the distance travelled by the mouse and is given by equation 1.

$$r_dist = \sum_{i=0}^{n-1} \sqrt{(posx_{i+1} - posx_i)^2 + (posy_{i+1} - posy_i)^2} \quad (1)$$

485 MOUSE ACCELERATION (MA) - The velocity of the mouse (in pixels/milliseconds) over the time (in milliseconds). A value of acceleration is computed for each interval defined by two consecutive MOUSE_UP and MOUSE_DOWN events, using the intervals and data computed for the Velocity.

490 TIME BETWEEN CLICKS (TBC) - the timespan between two consecutive MOUSE_UP and MOUSE_DOWN events, i.e., how long did it took the individual to perform another click.

DOUBLE CLICK DURATION (DCD) - the timespan between two consecutive
 495 MOUSE_UP events, whenever this timespan is inferior to 200 milliseconds. Wider timespans are not considered double clicks.

AVERAGE EXCESS OF DISTANCE (AED) - this feature measures the average excess of distance that the mouse travelled between each two consecutive
 500 MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*, respectively

in the coordinates $(x1, y1)$ and $(x2, y2)$. To compute this feature, first it is measured the distance in straight line between the coordinates of *mup* and *mdo* as $s_dist = \sqrt{(x2 - x1)^2 + (y2 - y1)^2}$. Then, it is measured the distance actually travelled by the mouse by summing the distance between each two consecutive MOUSE_MV events. Let us assume two vectors *posx* and *posy*, of size *n*, holding the coordinates of the consecutive MOUSE_MV events between *mup* and *mdo*. The distance actually travelled by the mouse, *real_dist* is given by equation 1. The average excess of distance between the two consecutive clicks is thus given by r_dist/s_dist .

AVERAGE DISTANCE OF THE MOUSE TO THE STRAIGHT LINE (ADMSL) - in a few words, this feature measures the average distance of the mouse to the straight line defined between two consecutive clicks. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$. Let us also assume two vectors *posx* and *posy*, of size *n*, holding the coordinates of the consecutive MOUSE_MOV events between *mup* and *mdo*. The sum of the distances between each position and the straight line defined by the points $(x1, y1)$ and $(x2, y2)$ is given by 2, in which *ptLineDist* returns the distance between the specified point and the closest point on the infinitely-extended line defined by $(x1, y1)$ and $(x2, y2)$. The average distance of the mouse to the straight line defined by two consecutive clicks is this given by s_dists/n .

$$s_dists = \sum_{i=0}^{n-1} ptLineDist(posx_i, posy_i) \quad (2)$$

DISTANCE OF THE MOUSE TO THE STRAIGHT LINE (DMSL) - this feature is similar to the previous one in the sense that it will compute the *s_dists* between two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*, according to equation 2. However, it returns this sum rather than the average value during the path.

530

SIGNED SUM OF ANGLES (SSA) - with this feature the aim is to determine if the movement of the mouse tends to "turn" more to the right or to the left. Let us assume three consecutive MOUSE_MOVE events, *mov1*, *mov2* and *mov3*, respectively in the coordinates $(x1, y1)$, $(x2, y2)$ and $(x3, y3)$. The angle α between the first line (defined by $(x1, y1)$ and $(x2, y2)$) and the second line (defined by $(x2, y2)$ and $(x3, y3)$) is given by $degree(x1, y1, x2, y2, x3, y3) = \tan(y3 - y2, x3 - x2) - \tan(y2 - y1, x2 - x1)$. Let us now assume two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*. Let us also assume two vectors *posx* and *posy*, of size n , holding the coordinates of the consecutive
540 MOUSE_MOVE events between *mup* and *mdo*. The signed sum of angles between these two clicks is given by equation 3.

$$rCls_angle = \sum_{i=0}^{n-2} degree(posx_i, posy_i, posx_{i+1}, posy_{i+1}, posx_{i+2}, posy_{i+2}) \quad (3)$$

ABSOLUTE SUM OF ANGLES (ASA) - this feature is very similar to the previous one. However, it seeks to find only how much the mouse "turned", independently
545 of the direction to which it turned. In that sense, the only difference is the use of the absolute of the value returned by function $degree(x1, y1, x2, y2, x3, y3)$, as depicted in equation 4.

$$rCls_angle = \sum_{i=0}^{n-2} | degree(posx_i, posy_i, posx_{i+1}, posy_{i+1}, posx_{i+2}, posy_{i+2}) | \quad (4)$$

DISTANCE BETWEEN CLICKS (DBC) - represents the total distance travelled
550 by the mouse between two consecutive clicks, i.e., between each two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, *mup* and *mdo*, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$. Let us also assume two vectors *posx* and

posy, of size n , holding the coordinates of the consecutive MOUSE_MOV events
555 between *mup* and *mdo*. The total distance travelled by the mouse is given by
equation 1.

4.3. Data Analysis

Prior to the training of the ANN we conducted an analysis of the data to
determine if there were statistically significant differences in the features when
560 comparing data collected in different levels of fatigue.

To achieve this we divided the data collected in two groups: one containing
the data collected while the users claimed to feel fatigued (values 1-3 of the
USAFAM scale) and the other containing the remaining of the data, collected
when users claimed to feel rested.

565 After this, we used the Mann-Whitney test to compare the distributions
of the values of each feature in order to verify that there were statistically
significant difference between the two groups of data (Table 2). The following
features have shown the most statistically significant differences: key down time,
mouse velocity, mouse acceleration, time between clicks, distance of the mouse
570 to the straight line, distance during clicks, average excess of distance and average
distance of the mouse to the straight line.

At this point, one of the concerns was to determine if the observed differences
were not due to natural variations of the use of the peripherals. To this end
we conducted an additional analysis, as follows. We took each group of data
575 (fatigued versus rested) and divided each group into two subgroups of equal
size, with the members of each group being selected randomly. We thus got
four groups, two containing data collected when individuals were fatigued and
other two when individuals were rested. We then used the Mann Whitney
test to determine if there were differences between the pairs of subgroups that
580 were extracted from the same group. We repeated this process 100 times and
computed the average value of the Mann Whitney test. Table 3 summarizes
these results. Only in the key down time feature were there found statistically
significant differences, in average, and only when comparing the two rested

Table 2: Analysis of significance of the collected data. Column a depicts the average value of the Mann Whitney test. Column b depicts the percentage of individuals for which the differences observed in each feature when fatigued and rested were statistically significant.

<i>Features</i>	<i>a</i>	<i>b</i>
Distance of the Mouse to the Straight Line	0.022	79%
Key Down Time	0.021	75%
Mouse acceleration	0.050	67%
Average Distance of the Mouse to the Straight Line	0.042	63%
Mouse velocity	0.023	63%
Average Excess of Distance	0.041	54%
Time Between Clicks	0.020	54%
Distance During Clicks	0.026	54%
Time between keys	0.271	46%
Total Excess of Distance	0.258	46%
Double Click Duration	0.003	33%
Absolute Sum of Angles	0.336	25%
Signed Sum of Angles	0.706	21%
Distance between clicks	0.0238	20%

subgroups. Given this, it is possible to conclude that the observed differences
585 can indeed be attributed to consistently different behaviours that are related to
the level of fatigue and not to natural variations.

However, and despite the observed differences, not every user revealed the
same behaviours or trends when under fatigue. The Average Distance of the
Pointer to the Line Between Clicks is an example of a metric in which the average
590 values increased in half of the users and decreased in the other half, despite the
statistically significant differences observed. This points out the importance of
developing individual models, trained with data from each user. It also points
out the subjectiveness of the effects of fatigue, which we had already previously
analysed in the detection of stress following similar approaches [45].

Table 3: Results of testing the intra-group differences by subdividing each group into two subgroups, randomly, and comparing them. Features are represented by their initials.

Features	average p -value (rested)	average p -value (tired)
KDT	0.086	0.024
MV	0.196	0.773
MA	0.211	0.610
AED	0.515	0.994
TBC	0.413	0.098
DMSL	0.669	0.531
DDC	0.983	0.095
ADMSL	0.539	0.787

595 Given this, and following similar approaches [31, 14], we centred the analysis in measures of variability, in an attempt to determine whether the variability of the two groups of data was different. We concluded that variability was significantly smaller in all features except for the Average Distance of the Mouse to the Straight Line (Figure 8) when users were rested, revealing that users
600 were more consistent in the use of the peripherals not fatigued (Table 4). As we can see, the differences in the values are much more consistent in the case of variability (i.e. nearly all the features follow an increasing tendency with the onset of fatigue) than in the measures of centrality. Hence the decision of using variability for training the neural network.

605 These differences are also graphically visible. Figures 4, 5, 6, 7 and 8 depict how the differences in the distributions of the variance (rightmost images) are larger than the differences in the distributions of the average values.

5. Real-Time Classification of Fatigue

The first step in the process of training a suitable classifier for the detection of
610 fatigue from the data of user interaction was to select the features that presented the most significant differences when comparing the two groups of data. The

Table 4: The values of centrality and variance for each feature, represented by their initials, and each group of data. The three last columns represent the following trends: A - Trend of the medians when comparing rested with fatigued; B - Trend of the means when comparing rested with fatigued; C - Trend of the variance when comparing rested with fatigued. In these columns, a value of +70% should be interpreted as "increases in 70% of the cases".

<i>Features</i>	<i>State</i>	<i>Variance</i>	<i>Median</i>	<i>Mean</i>	<i>A</i>	<i>B</i>	<i>C</i>
KDT	Rested	1177.6	121.8	139.1	+ 100%	+ 50%	+ 100%
	Fatigued	4994.2	116.8	138.8			
MV	Rested	0.3710	0.182	0.191	+ 75%	+ 50%	+ 100%
	Fatigued	0.7344	0.188	0.199			
MA	Rested	0.4128	0.740	0.801	+ 60%	- 60%	+ 90%
	Fatigued	0.7520	0.786	0.844			
TBC	Rested	1.3342	3539.0	3925.0	- 51%	- 51%	+ 100%
	Fatigued	3.2017	2794.0	2913.0			
DMSL	Rested	62620.7	4232.0	4047.0	- 50%	- 59%	+ 77%
	Fatigued	39489.9	3293.0	4366.0			
DDC	Rested	28233.7	116.3	148.00	+ 77%	+ 67%	+ 77%
	Fatigued	55839.0	87.99	93.5			
AED	Rested	1.8564	2.080	1.991	+ 50%	+ 60%	+ 90%
	Fatigued	5.9635	1.743	1.818			
ADMSL	Rested	7765.5	28.100	29.5	- 60%	- 50%	- 60%
	Fatigued	4311.1	27.040	30.5			

selected inputs of the neural network were thus the following features, supported by the data analysis detailed in the previous section: key down time, velocity mouse, mouse acceleration, time between clicks, distance from pointer to line
615 between clicks, clicks during distance, average excess of distance and average distance of the mouse to the straight line and also the part of the day (because of the influence of the circadian rhythm). The output of the artificial neural network is a value between 1 and 7 that denotes the seriousness of the level of fatigue, similarly to the USAFAM questionnaire.

620 With the input and output layers defined, a multilayer feed-forward neural network was used to model the relationship between behavioural features and

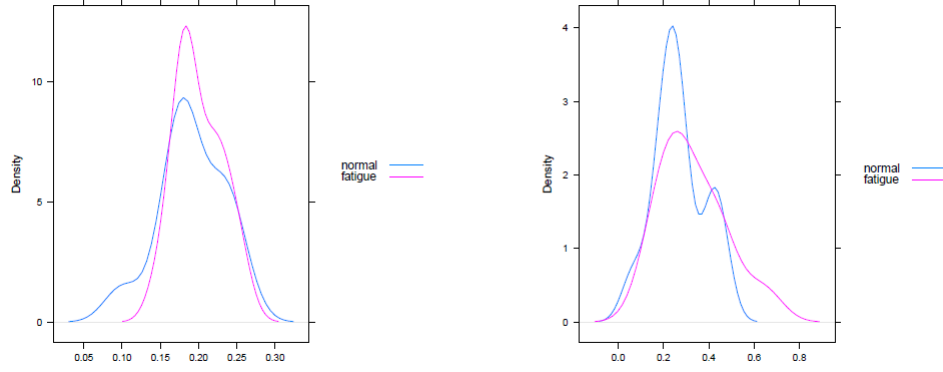


Figure 4: Differences in the distribution of the means (left image) and in the distribution of the variance (right image) for the feature *Mouse Velocity*.

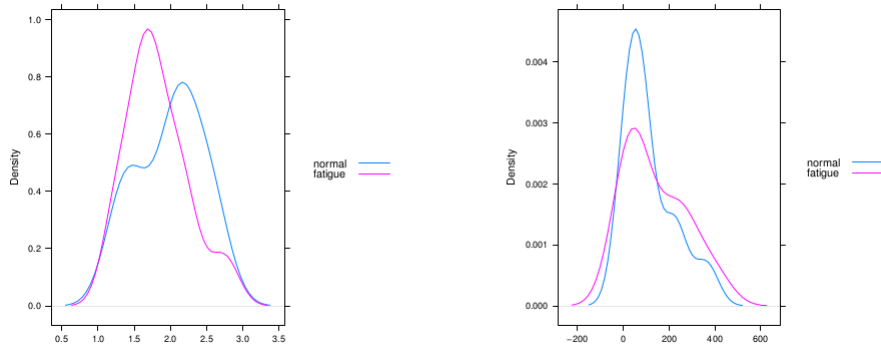


Figure 5: Differences in the distribution of the means (left image) and in the distribution of the variance (right image) for the feature *Average Excess of Distance*.

a level of fatigue. This is a supervised network, where the network knows the outputs, and through these attempts to adjust the weights of the initial and intermediate nodes in order to get results close to the expected ones. The back-
625 propagation learning algorithm was used that, besides the input and output layer, uses a intermediate hidden layer that lies between the input layer and the output layer, with a total of 10 nodes, as can be seen in Figure 9. The number of hidden nodes has been selected after the carrying out of several performance tests with different numbers of hidden nodes.

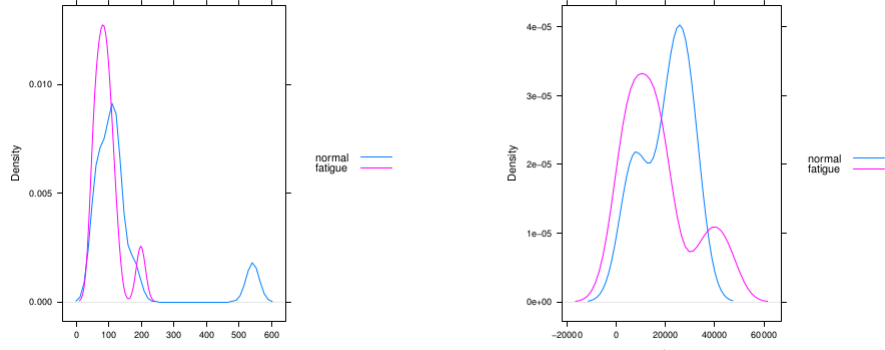


Figure 6: Differences in the distribution of the means (left image) and in the distribution of the variance (right image) for the feature *Distance during clicks*.

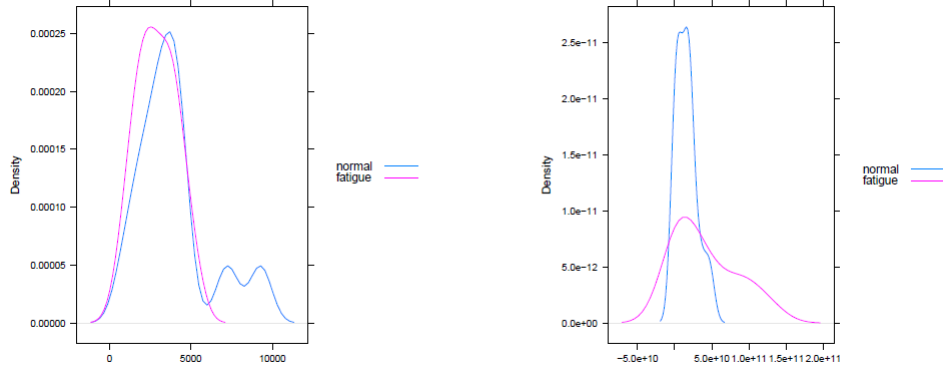


Figure 7: Differences in the distribution of the means (left image) and in the distribution of the variance (right image) for the feature *Time between clicks*.

630 The network was trained with a dataset containing a total of 74 instances, each of which containing the value of the variance of each of the features over a period of one hour and the level of subjective fatigue provided by the users through the questionnaire, for that period. The network was trained during 125 iterations where it reached a minimum RMSE (Root-Mean-Square Error)
 635 as can be seen in Figure 10.

The relative importance of the variables used as input was assessed through a method proposed by Garson [46]. This method identifies the relative importance

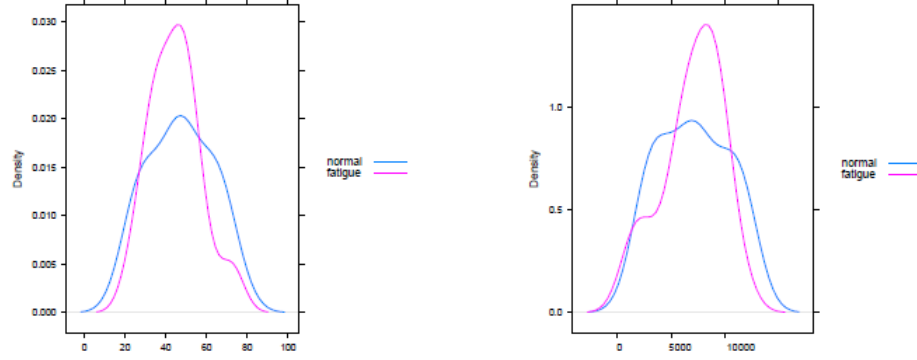


Figure 8: Differences in the distribution of the means (left image) and in the distribution of the variance (right image) for the feature *Average Distance of the Mouse to the Straight Line*.

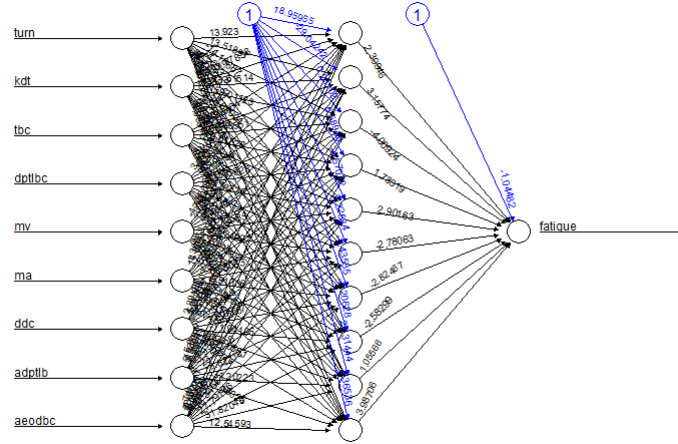


Figure 9: Structure of the the ANN trained. This is a three-layered feed-forward neural network.

of input variables to the final output on supervised neural network using the weights of each node of the model. The relative importance (or the strength of association) of an input variable to the output can be determined by identifying all connections between nodes' weighted interest. That is, all the weights of specific binding of the input node which pass through to the hidden layer output variable are identified. This process is repeated for all other input variables

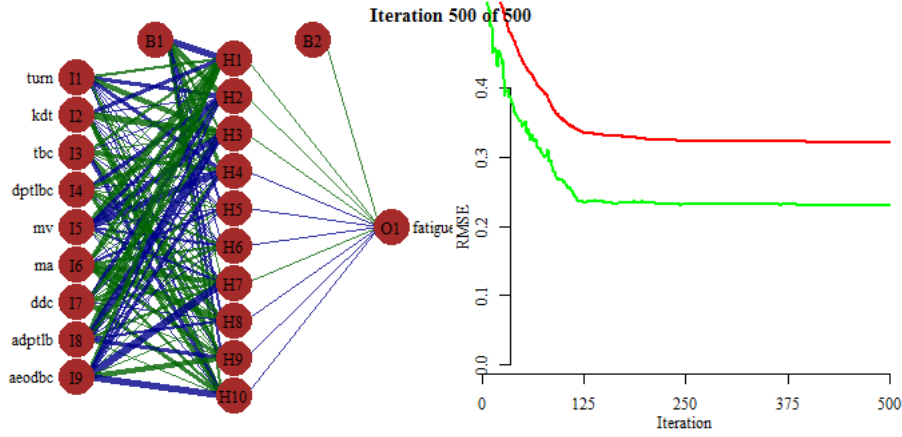


Figure 10: Evolution of the RMSE during the evolution of the training of the ANN (green line) and during validation (red line).

to obtain a list of all weights which are specific to each input variable. The connections are computed for each node of the input layer and balanced against all other entries. A single value is obtained for each input variable that describes the relationship with the model output.

Figure 11 shows that x5 and x6 (mouse velocity and acceleration, respectively) have the strongest positive relationships, while the time between clicks and key down time have more negative relationships with the output. It is also visible that the shift/turn (morning, afternoon, evening) in which the data was collected has little influence on the output, which indicates that the time of the day is less important than the other features when it comes to classify fatigue.

5.1. Results

The ANN trained and outlined above was tested and validated with data from the second week of the data collection period, in which each instance, as described above, contained the variance of each feature in the period of an hour and the subjective measure of the level of fatigue of the user. This approach allows us to compare the value pointed out by the user against the value provided as output by the ANN. The main result is that the trained ANN

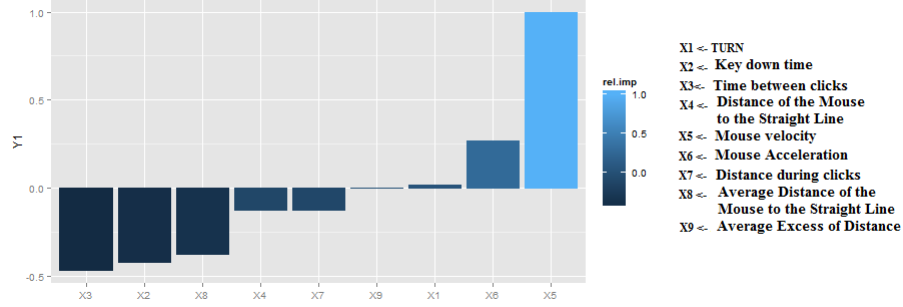


Figure 11: Relative importance of different variables.

correctly classified 81% of the instances, i.e., producing as output the same value that the user provided in the questionnaire. The 19% that were misclassified where nonetheless classified as neighbouring values, as can be seen in Table 5.

Table 5: Results of the validation of the ANN. 81% of the instances were correctly classified (green cells). The 19% misclassified instances where nonetheless classified as neighbouring values (red cells).

		PREDICTED						
		1	2	3	4	5	6	7
U S A F S A M	1	9	2	0	0	0	0	0
	2	3	13	3	0	0	0	0
	3	0	2	15	0	0	0	0
	4	0	0	1	8	1	0	0
	5	0	0	0	0	8	1	0
	6	0	0	0	0	1	6	0
	7	0	0	0	0	0	0	1

6. Discussion

665 We have until now described the main results of this research line. Let us now provide some more insights that can better allow to understand and interpret these results. First of all, it is important to point out that our main aim

is to develop a fatigue monitoring tool that can be used in real-time, without interfering in any way with the worker’s routine. To this end, we developed an
670 approach based on the monitoring of the user’s performance. The main conclusion, as expected, is that when the users claim to be fatigued, their performance decreases. This is evidenced by less efficient interaction patterns (e.g. longer mouse clicks, larger distances travelled by the mouse, longer key down times). This conclusion, which comes as no surprise, is nonetheless important in the
675 sense that it allows, for the first time, to quantify and study these differences.

As analysed previously, there are some potential drawbacks associated to this kind of performance-based approaches, namely the difficulty in precisely accounting for changes in performance (e.g. they may not be entirely due to mental fatigue). Moreover, measuring performance is not always easy. The type
680 of task being performed may significantly influence the interaction patterns. The workload is another issue to consider. Task standardization, which would to some extent solve these drawbacks, may be difficult to achieve. The motivation of the worker or feelings of boredom will also influence performance.

To address these issues, we are also working towards classifying the type of
685 task (e.g. text typing, web browsing) from the interaction patterns, identifying workload and quantifying the level of task attention (Pimenta et al., 2014). This kind of information, which to some extent describes the user’s context, will enable the development of more accurate classifiers.

Nonetheless, and despite these efforts, the accuracy of the approach will,
690 evidently, never match that of other approaches such as the ones based on physiological sensors. But then again, that was not the main aim from the beginning.

The main aim was to develop an approach that could be used realistically in a common workplace environment, especially in those workplaces in which
695 workers spend long hours using a computer. To this end, this research line accomplishes its objectives, as the results detailed above show.

Given that different effects due to fatigue could be found in different users, we are now working towards an increased personalization, in which each user has

an ANN trained with their own data. Moreover, we are also working towards
700 the validation of the ANN using Electroencephalography in parallel with our
approach, through a cooperation with researchers of the University of Gronin-
gen.

7. Conclusion

Fatigue can nowadays be seen as a major health issue. Mental fatigue is
705 especially on the rise given the dramatic change in the labour market, from
mostly physical jobs to very mentally demanding ones. Harsher work conditions,
increased competitiveness or high unemployment add to the pressure of today's
worker and amplify the negative effects of mental fatigue. These act on several
dimensions, including the personal, social, emotional, organizational, among
710 others. In a general way, the need for better strategies for fatigue management
is nowadays acknowledged.

The main problem is that existing approaches are unsuited to be used re-
alistically in nowadays workplaces, mostly because they are either invasive, ex-
pensive, require specialized personnel or interfere with the worker's routines.

715 In this work we propose an approach that is suited for such purpose. Indeed,
it can be used continuously to quantify the level of fatigue of a worker without
the need for conscious or explicit actions by the worker. This approach is based
on the worker's interaction patterns with the computer when using the mouse
and the keyboard. Indeed, in this work we show that when users claim to feel
720 fatigued, they use these peripherals significantly differently.

The central part of the work was thus to train an ANN that could classify
the level of fatigue taking as input the interaction patterns. To this end, we
conducted a study with the duration of two weeks in which the users of com-
puters, with their interactions being constantly monitored, would also regularly
725 provide a subjective measure of their level of fatigue. This allowed to verify that
the usage of the computer significantly varies with the level of fatigue. Having
established this, we moved on to train and test an ANN that successfully classi-

fied the level of fatigue of the users in the following week. This was achieved by checking the value of fatigue provided by the ANN against the subjective value
730 provided by the participants.

Given this, we can conclude that this approach can be considered as a valid one in the panorama of fatigue monitoring and detection systems. Its main advantages are that it is inexpensive (it relies on existing infrastructures), it is easy to use (requires only the installation of some software) and can be used
735 continuously throughout the workday in a non-intrusive way and without interfering with the worker's routines. Ultimately, these approaches will improve access to fatigue management initiatives, with a positive impact on healthcare, productivity, work quality and other indicators.

8. Acknowledgements

740 This work is part-funded by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within projects FCOMP-01-0124-FEDER-028980 (PTDC/EEISII/1386/ 2012) and PEst-OE/
745 EEI/UI0752/2014.

References

- [1] M. M. Mitler, J. C. Miller, Some practical considerations and policy implications of studies of sleep patterns, *Behavioral Medicine* 21 (4) (1996) 184–185.
- 750 [2] J. C. Miller, Fundamentals of shiftwork scheduling, *Ergonomics in Design: The Quarterly of Human Factors Applications* 16 (3) (2008) 13–17.
- [3] S. Folkard, P. Tucker, Shift work, safety and productivity, *Occupational medicine* 53 (2) (2003) 95–101.

- [4] D. van der Linden, P. Eling, Mental fatigue disturbs local processing more
 755 than global processing., *Psychological research* 70 (5) (2006) 395–402. doi :
 10.1007/s00426-005-0228-7.
 URL <http://www.ncbi.nlm.nih.gov/pubmed/15968553>
- [5] R. J. Williamson, S. Purcell, A. Sterne, S. Wessely, M. Hotopf, A. Farmer,
 P. C. Sham, The relationship of fatigue to mental and physical health in a
 760 community sample., *Social psychiatry and psychiatric epidemiology* 40 (2)
 (2005) 126–32. doi:10.1007/s00127-005-0858-5.
 URL <http://www.ncbi.nlm.nih.gov/pubmed/15685404>
- [6] L. Perelli, Fatigue Stressors in Simulated Long-Duration Flight. Effects on
 Performance, Information Processing, Subjective Fatigue, and Physiologi-
 765 cal Cost (March 1977).
 URL [http://oai.dtic.mil/oai/oai?verb=](http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA105484)
[getRecord&metadataPrefix=html&identifier=ADA105484](http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA105484)
- [7] M. H. Bonnet, Sleep, performance and mood after the energy-expenditure
 equivalent of 40 hours of sleep deprivation, *Psychophysiology* 17 (1) (1980)
 770 56–63.
- [8] J. C. Miller, In Search of Circasemidian Rhythms, Tech. rep., Brooks City-
 Base TX (Nov. 2006).
- [9] R. Hockey, Stress and fatigue in human performance, Vol. 3, John Wiley
 & Sons Inc, 1983.
- 775 [10] M. M. Lorist, M. Klein, S. Nieuwenhuis, R. Jong, G. Mulder, T. F. Mei-
 jman, Mental fatigue and task control: planning and preparation, *Psy-
 chophysiology* 37 (5) (2000) 614–625.
- [11] A. Pimenta, D. Carneiro, P. Novais, J. Neves, Monitoring Mental Fatigue
 through the Analysis of Keyboard and Mouse Interaction Patterns, *Hybrid*
 780 *Artificial Intelligent ...* (2013) 222–231.

URL http://link.springer.com/chapter/10.1007/978-3-642-40846-5_23

[12] A. T. Welford, Fundamentals of skill.

[13] S. Folkard, Shift work, safety and productivity, *Occupational Medicine* 53 (2) 95–101. doi:10.1093/occmed/kqg047.

785

URL <http://occmed.oupjournals.org/cgi/doi/10.1093/occmed/kqg047>

[14] C. K. McCleron, J. C. Miller, Variance as a Measure of Performance in an Aviation Context, *The International Journal of Aviation Psychology* 21 (4) (2011) 397–412. doi:10.1080/10508414.2011.606765.

790

URL <http://www.tandfonline.com/doi/abs/10.1080/10508414.2011.606765>

[15] S. G. Hart, L. E. Staveland, Development of nasa-tlx (task load index): Results of empirical and theoretical research, *Advances in psychology* 52 (1988) 139–183.

795

[16] G. B. Reid, F. T. Eggemeier, C. A. Shingledecker, Subjective workload assessment technique, Tech. rep., DTIC Document (1982).

[17] L. Ames, E. George, Revision and Verification of a Seven-Point Workload Estimate Scale.

800

URL <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA269194>

[18] D. McNair, M. Lorr, L. Droppleman, Profile of mood states (POMS).

URL <http://ubir.buffalo.edu/xmlui/handle/10477/1888>

[19] D. McNair, M. Lorr, L. Droppleman, Manual for the profile of mood states, 1971, Educational and Industrial Testing Service, San Diego.

805

[20] T. Åkerstedt, M. Gillberg, Subjective and objective sleepiness in the active individual, *International Journal of ...* 52 (1-2) (1990) 29–37.

URL <http://informahealthcare.com/doi/abs/10.3109/00207459008994241>

- 810 [21] M. Gillberg, G. Kecklund, T. Å kerstedt, Relations between performance and subjective ratings of sleepiness during a night awake., *Sleep: Journal of Sleep* ...17 (3) (1994) 236–241.

URL <http://psycnet.apa.org/psycinfo/1995-03957-001>

- [22] E. Hoddes, V. Zarcone, H. Smythe, Quantification of sleepiness: a new approach, ...10 (4) (1973) 431–436.

815 URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1469-8986.1973.tb00801.x/full>

- [23] a. Gevins, H. Leong, R. Du, M. E. Smith, J. Le, D. DuRousseau, J. Zhang, J. Libove, Towards measurement of brain function in operational environments., *Biological psychology* 40 (1-2) (1995) 169–86.

820 URL <http://www.ncbi.nlm.nih.gov/pubmed/7647178>

- [24] A. Gevins, M. Smith, Detecting transient cognitive impairment with EEG pattern recognition methods, *Aviation, space, and environmental medicine*.

URL <http://www.eeg.com/DetectingTransientCognitive.pdf>

- 825 [25] R. Goldstein, L. O. Bauer, J. A. Stern, Effect of task difficulty and inter-stimulus interval on blink parameters, *International Journal of Psychophysiology* 13 (2) (1992) 111–117. doi:10.1016/0167-8760(92)90050-L.

URL <http://www.sciencedirect.com/science/article/pii/016787609290050L>

- 830 [26] J. A. Stern, L. C. Walrath, R. Goldstein, The Endogenous Eyeblick, *Psychophysiology* 21 (1) (1984) 22–33. doi:10.1111/j.1469-8986.1984.tb02312.x.

URL <http://doi.wiley.com/10.1111/j.1469-8986.1984.tb02312.x>

- [27] L. N. Orchard, J. a. Stern, Blinks as an index of cognitive activity during reading., *Integrative physiological and behavioral science : the official jour-*

835

nal of the Pavlovian Society 26 (2) 108–16.

URL <http://www.ncbi.nlm.nih.gov/pubmed/1878317>

- [28] G. JEAN-LOUIS, H. V. GIZYCKI, F. ZIZI, A. SPIELMAN, P. HAURI,
H. TAUB, THE ACTIGRAPH DATA ANALYSIS SOFTWARE: I. A
NOVEL APPROACH TO SCORING AND INTERPRETING SLEEP-
WAKE ACTIVITY, *Perceptual and Motor Skills* 85 (1) (1997) 207–216.
doi:10.2466/pms.1997.85.1.207.
URL <http://www.amsci epub.com/doi/abs/10.2466/pms.1997.85.1.207?journalCode=pms>
- [29] N. K. Tang, A. G. Harvey, Correcting distorted perception of sleep in
insomnia: a novel behavioural experiment?, *Behaviour Research and
Therapy* 42 (1) (2004) 27–39. doi:10.1016/S0005-7967(03)00068-8.
URL <http://linkinghub.elsevier.com/retrieve/pii/S0005796703000688>
- [30] M. Stork, Some Methods Systems and Sensors which are Possible for
Driver’s Drowsiness Estimation, *wseas.us* 87–94.
URL <http://www.wseas.us/e-library/conferences/2012/Barcelona/CSCS/CSCS-14.pdf>
- [31] S. C. Segerstrom, L. S. Nes, Heart rate variability reflects self-regulatory
strength, effort, and fatigue., *Psychological science* 18 (3) (2007) 275–81.
doi:10.1111/j.1467-9280.2007.01888.x.
URL <http://pss.sagepub.com/content/18/3/275.short>
- [32] C. Wickens, J. McCarley, *Applied attention theory*, CRC Press, Boca
Raton FL, 2007.
URL http://www.prometei.de/fileadmin/prometei.de/veranstaltungen/2008-05-28-Wickens_AppliedAttentionTheory.pdf
- [33] R. Parasuraman, G. Wilson, Putting the brain to work: Neuroergonomics
past, present, and future, *Human Factors: The Journal of the ...* 50 (3)

(2008) 468–474.

865 URL <http://hfs.sagepub.com/content/50/3/468.short>

- [34] C. Spielberger, R. Gorsuch, R. Lushene, Manual for the state-trait anxiety inventory.

URL <http://ubir.buffalo.edu/xmlui/handle/10477/2895>

- [35] C. Spielberger, R. Gorsuch, R. Lushene, Self-Evaluation Questionnaire: STAI Form X-1.

870 URL <http://ubir.buffalo.edu/xmlui/handle/10477/1878>

- [36] S. Curran, Short Form of the Profile of Mood States (POMS-SF): Psychometric information., *Psychological Assessment* 7 (1) (1995) 80–83. doi: 10.1037//1040-3590.7.1.80.

875 URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/1040-3590.7.1.80>
<http://psycnet.apa.org/journals/pas/7/1/80/>

- [37] K. Millar, M. Jellicic, B. Bonke, a. J. Asbury, Assessment of preoperative anxiety: comparison of measures in patients awaiting surgery for breast cancer., *British journal of anaesthesia* 74 (2) (1995) 180–3.

880 URL <http://www.ncbi.nlm.nih.gov/pubmed/7696068>

- [38] P. Newcombe, G. Boyle, High school students’ sports personalities: Variations across participation level, gender, type of sport, and success, *International Journal of Sport ...* (1995) 277–294.

URL http://epublications.bond.edu.au/hss_pubs/802/

- 885 [39] J. Casali, W. Wierwille, A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications, *Human Factors: The Journal of the ...* 25 (6) (1983) 623–641.

URL <http://hfs.sagepub.com/content/25/6/623.short>

- 890 [40] W. Wierwille, J. Gutmann, Secondary task measurement of workload as a function of simulated vehicle dynamics and driving conditions, *Human*

Factors: The ... 19 (6) (1977) 557–565.

URL <http://hfs.sagepub.com/content/19/6/557.short>

- 895 [41] C. Ramos, J. C. Augusto, D. Shapiro, Ambient intelligence—the next step
for artificial intelligence, *Intelligent Systems, IEEE* 23 (2) (2008) 15–18.
- [42] J. Agüero, M. Rebollo, C. Carrascosa, V. Julián, Mdd-approach for de-
veloping pervasive systems based on service-oriented multi-agent systems,
ADCAIJ: Advances in Distributed Computing and Artificial Intelligence
Journal 1 (6) (2013) 55–64.
- 900 [43] P. Giner, V. Pelechano, An architecture to automate ambient business sys-
tem development, in: *Ambient Intelligence*, Springer, 2008, pp. 240–257.
- [44] A. Woodrow, Cognitive Performance Research at Brooks Air Force Base,
Texas, 1960-2009, Aviation, Space, and Environmental ...
- [45] D. Carneiro, P. Novais, J. Neves, Conflict Resolution and its Context: From
905 the Analysis of Behavioural Patterns to Efficient Decision-Making, Vol. 18,
Springer, 2014.
- [46] G. D. Garson, Interpreting neural-network connection weights, *AI expert*
6 (4) (1991) 46–51.